

DOCUMENT RESUME

ED 307 697

EA 021 021

AUTHOR Folger, John
TITLE Project STAR: Design and Measurement.
PUB DATE Mar 89
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989). For related document, see EA 021 020.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Class Size; Educational Legislation; Primary Education; Program Design; *Program Development; *School Districts; *Small Classes; Teacher Student Ratio
IDENTIFIERS *Project STAR; Tennessee

ABSTRACT

The Tennessee legislature sponsored a study on the effects of class size (Project STAR). The design and measurement techniques of this study are discussed. The establishment of the basic elements of the design is examined, followed by the recruitment of schools and school systems to participate in the study. The process of setting up smaller classes involved student achievement measures, teacher training, and selection of class type in order to provide objective cumulative effects. Appended are three references and a table of the basic design of Project STAR. (SI)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED307697

AERA - 1989

PROJECT STAR, DESIGN AND MEASUREMENT
John Folger

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
DERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

John Folger

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper presented at the Annual Meeting of the American Educational
Research Association, San Francisco, California, March 27-31, 1989

BEST COPY AVAILABLE

EA 021 021

PROJECT STAR, DESIGN AND MEASUREMENT
John Folger

The sponsor of the Tennessee legislation to study class size, Representative Steve Cobb, was social scientist (Ph.D. in Sociology) and lawyer as well as a legislator. He wanted a definitive study of the question of the effects of class size, one that was well designed, and conducted in a manner that would assure that its findings were accurate and credible. The legislation specified several aspects of the design, suggested others that were optional, and provided ample (3 million dollars per year for four years) resources to implement the design.

The State Department contracted with the four universities (Memphis State, Tennessee State, University of Tennessee, and Vanderbilt) to provide technical assistance in the design and conduct of the study. When the universities joined the project, there were about two months to establish the basic elements of the design, recruit school systems and schools to participate in the study, and get the smaller classes set up.

An extensive literature review had been completed the previous year, by a professor at Memphis State, which provided some suggestions about factors that should be included in a design. The class size literature does not provide a clear conceptual framework about how class size effects are achieved, and therefore provides limited guidance about variables that should be included in a study design. The literature on effective schools provides more guidance about variables that make a difference (for example: Brophy, 1986; Good, 1986; Walberg, 1984). More guidance was provided by the

legislation itself, which specified a number of design characteristics. These include the size of small classes. (They were to average 15 students, and could range from 13 to 17.) They were to be compared with regular classes. (Tennessee K through 3 grade average class enrollment is about 23 to 1 and a range of 22 to 25 was establish.) Participating schools were to be selected from east, middle and west Tennessee, and from "inner-city, suburban, urban, and rural schools." Teachers were to be trained. Student achievement was to be the primary criterion for judging the effectiveness of class size reductions. Other outcomes (teacher morale, pupil attendance) could also be included but were not mandated. Cost effectiveness of class size reduction was to be estimated.

In meetings to plan the study, State Department of Education and the legislative sponsors added the assignment of a full-time teacher aide to regular-sized classes as another design requirement. The university research participants took the legislation and other design criteria and developed a within-school design, with random assignment of teachers and pupils to the specific class type. The design is outlined in Figure One.

The three class types were: (1) small classes with an enrollment of 13-17; (2) regular classes, with enrollment of 22-25 who serve as a primary control group; and (3) regular classes (22-25) with a full-time aide. To participate, a school had to have a total kindergarten enrollment of at least 57, the minimum enrollment to set up three classes. If enrollment exceeded 67, an additional class or classes had to be established. Participant schools had a minimum of three classes, and a maximum of eleven, although only one school had more than 7. This produced a design that was not balanced across class types in every school.

All 140 school systems within the state were invited by the Commissioner of Education to participate in the project; about 55 systems volunteered. Elementary schools that had all grades, K-3, and met the minimum enrollment criterion of 57 were invited to participate. There were about 110 schools that volunteered that met the criterion. From this total we chose schools to provide a regional balance, and to provide a balanced number of classes in each type of school (inner-city, suburban, urban or rural). A power test was conducted to determine what sample size was needed to detect a difference of 3% between class types at the .01 level of confidence. This indicated that a sample of 70 to 80 classes of each type was needed. We decided to over sample to protect against schools that might drop out or other forms of attrition which might occur over a four-year period. The initial goal was to select a sample with at least 100 classes of each type, or a total of 300.

The first year there were 328 classes: 128 small classes, 101 regular classes, and 99 classes with an aide. The second year (first grade) there were 128 small classes, 110 regular classes, and 108 regular/aide classes. In the third year there were 131 small classes, 102 regular classes, and 107 regular/aide classes.

Comparison Sample of Schools

A comparison group of 22 schools with 57 teachers and about 1300 pupils was recruited from 13 of the 42 school systems that had schools involved with Project STAR. The comparison schools allowed us to check for the "Hawthorne effect" and to collect comparable test score data which is not available from statewide testing for all grades. Selection of comparison schools was based on similarity with the project schools in terms of type of students being served. In these schools, the same tests were administered, and the testing

was monitored in the same way to try to assure comparability. There were no significant differences in test scores in kindergarten between the STAR schools and the comparison schools in the same school systems. The comparison school sample had no inner-city schools because we could not recruit any from the participating school systems.

Process Measures

Most of the studies of class size have focused on whether or not there was an effect on student achievement, without identifying how teachers changed their teaching when they had fewer students. The staff planning Project STAR wanted to include information on how teachers changed their teaching when they had a small class or when they had an aide. A sample of about 85 teachers was observed between one and three times a year in the second grade, and another sample was observed in the third grade. All teachers in the project completed a log indicating how they spent their time during a "typical" day. Logs were completed three times in kindergarten, once in first and second grades, and twice in the third grade. Each year teachers were interviewed near the end of the school year to find out how they had modified their teaching. All teachers completed a teacher problem checklist to rate the problems they encountered. A special study is being made of teachers whose pupils had large achievement gain. It involves both observation and interviews with successful teachers. From these various sources it is possible to identify some of the changes in teaching behavior that occur when teachers have an aide or have a small class.

Measures of Achievement

Students were tested each year with the Stanford Achievement Test. In the primary grades the major emphasis of those tests is on reading, language

arts and math. We compared the Stanford test item coverage with the learning objectives in the Tennessee Basic Skills First (BSF) curriculum. BSF is a common set of reading and math learning objectives that all schools are supposed to teach in the elementary grades. Because there was only about a 60-70% overlap between the Stanford Test items and the curriculum that was being taught, the State Department of Education developed a criterion test keyed specifically to the BSF objectives in reading and math. New tests were prepared for the first and second graders; there was already a similar state criterion test developed for the third grade.

All of the testing of pupils was monitored by trained substitute teachers. Teachers administered the test, monitors assisted with the work of distributing and collecting the test booklets, and observed all of the testing in the school to be sure that standard procedures were followed in the testing.

Training Design

For second and third grade teachers, a three-day before-school training program was designed and implemented. A random sample of 13 of the 75 STAR schools were selected. Of the 58 second grade teachers in the 13 schools, 52 were trained; the other six were unavailable at the time the training program was offered. Trained teachers had five follow-up sessions during school to assist them in implementing what they had learned in the training sessions. Teachers in the training program were observed in the spring of the year before they were trained, and were observed twice after training. The design permits comparison of student growth in achievement of the trained teachers compared with all teachers who were not trained. The training sample is not large enough to permit comparisons between training and no training by school

type. A sample of 32 teachers in eight schools that were not involved in the training were also observed. The training program is being repeated in the same 13 schools in the third grade, but there were some changes in the instructional design. The first year all second grade teachers received the same training curriculum, and they found out about their class assignment after training was complete. The second year the third grade teachers got their class assignments (small, regular, regular/aide) before training, and one day of training was focused specifically on preparation for the kind of class in which they were going to be working.

Cumulative Effects of Class Type

One design objective specified in the legislation was that the project should measure the cumulative effects of being in a small class over the four-year period of the project. To achieve this objective, students were initially randomly assigned to a class type in kindergarten, and continued in that class type for four years, with one exception. Between kindergarten and first grade, students could be reassigned between regular and regular/aide classes but students in small classes remained in small classes. New students entering the school each year were randomly assigned to replace dropouts.

Sometimes growth or shrinkage of enrollment in a school required that a class be added or dropped and that usually led to some reassignment of pupils from one class type to another.

Some project schools had fairly stable enrollment and very few dropouts or additions, while others (most often in metropolitan areas had over 50% dropouts and additions.

The design permits comparison of class type effects for students who have only been in the project at one grade level (first, second or third) with

students who have been in the project every year. The inclusion of more classes in the project than were necessary to study single-year effects will enable us to study longitudinal effects even where attrition has reduced the original cohort to one-half to one-third of its original size. Since the class is the primary unit of analysis, class means can be calculated even when remaining students of the original cohort are less than half of the total.

The design selected was reviewed by outside consultants, one of whom, Dr. Jeremy Finn, has also participated in the analysis which has gotten more complex each year as additional data is collected. In general, the design has proven adequate to provide useful and definitive tests of the original questions that the legislation sought answers for.

REFERENCES

Brophy, Joe "Teacher Influences on Student Achievement" American Psychologist
Vol. 41, Oct. 1986, pp. 1069-1077.

Good, Tom and Ronda Weinstein "Schools Make a Difference" American Psychologist Vol. 41, October 1986, pp. 1090-97.

Walberg, Herbert "Improving the Productivity of America's Schools" Educational Leadership (May 1984), pp. 19-29.

Figure One

Basic Design of Project STAR
With Number of Classes in Each Cell of the Design
Second Grade, 1987-88

School Type and Training	Class Type (Each school has at least 3 classes.)			Total
	Small	Regular	Regular/Aide	
Inner City				
Training	3	4	3	10
No Training	23	19	20	62
Suburban				
Training	4	5	4	13
No Training	23	16	18	57
Rural				
Training	14	9	9	32
No Training	50	39	40	129
Urban				
Training	1	1	1	3
No Training	13	9	12	34
TOTAL				
Training	22	19	17	58
No Training	109	83	90	282

There was no training program in kindergarten and first grade. As a secondary control group, there were 21 "comparison" schools (with 57 classes) drawn from 13 of the 42 school districts that had STAR schools. All classes in the comparison schools were regular classes; only student achievement data was collected from them. The primary control group was the regular classes within the Project STAR schools.